Behavior in Repeated Prisoner's Dilemma Games with Shifted Outcomes Analyzed with a Statistical Learning Model

Van Assen, Marcel; Snijders, Chris; Weesie, Jeroen

Published in:
The Journal of Mathematical Sociology, 30:2, 159 - 180

DOI:
10.1080/00222500600549670

The Journal of Mathematical Sociology
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713618269

Behavior in Repeated Prisoner's Dilemma Games with Shifted Outcomes Analyzed with a Statistical Learning Model

Marcel Van Assen [a]; Chris Snijders [b]; Jeroen Weesie [c]
[a] Tilburg University, The Netherlands
[b] Eindhoven University, The Netherlands
[c] Utrecht University, The Netherlands

Online Publication Date: 01 April 2006
To cite this Article: Van Assen, Marcel, Snijders, Chris and Weesie, Jeroen (2006) 'Behavior in Repeated Prisoner's Dilemma Games with Shifted Outcomes Analyzed with a Statistical Learning Model', The Journal of Mathematical Sociology, 30:2, 159 - 180
To link to this article: DOI: 10.1080/00222500600549670
URL: http://dx.doi.org/10.1080/00222500600549670

Journal of Mathematical Sociology, 30:159–180, 2006

ISSN: 0022-250X print/1545-5874 online
DOI: 10.1080/00222500600549670

Routledge
Taylor & Francis Group

Behavior in Repeated Prisoner's Dilemma Games with Shifted Outcomes Analyzed with a Statistical Learning Model

Marcel van Assen
Tilburg University, The Netherlands

Chris Snijders
Eindhoven University, The Netherlands

Jeroen Weesie
Utrecht University, The Netherlands

We address whether cooperative behavior in a repeated Prisoner's Dilemma (PD) is more easily achieved under "good circumstances" (all payoffs in the constituent PD are positive), "bad circumstances" (payoffs are negative), or "mixed circumstances." To analyze the behavior in these repeated PDs, we developed and applied a learning model that improves upon standard learning models in two ways: (1) It allows for statistical tests of the parameter estimates, and (2) it allows for the incorporation of independent variables (e.g., subject or game characteristics). The model is applied to the data of the repeated PD experiment in van Assen and Snijders (2004, 2005). Our findings demonstrate that our model can be used to identify and test how learning differs across persons and across different circumstances.

Keywords: prisoner's dilemma, collective action, learning, learning models, risk, risk preferences

INTRODUCTION

We focus on actors' behavior in (repeated) prisoners' dilemmas that are similar except for a shift in their outcomes. More specifically, the research question investigated in the present paper is whether actors differ in their inclination to cooperate in "negative," "mixed,"

Address correspondence to Marcel van Assen, Department of Methodology and Statistics (MTO), Tilburg University, Warandelaan 2, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: M.A.L.M.vanAssen@uvt.nl

and "positive" prisoners' dilemmas. Negative prisoners' dilemmas have only negative outcomes in the constituent game, positive dilemmas have only positive outcomes, while the best outcomes in mixed dilemmas are positive and the worst are negative. In order to answer this research question a new methodology, a statistical learning model, is developed in the present paper. After describing the model and its advantages compared with traditional methodology, the research question is answered by applying the model to an existing data set.

The question whether actors are more prone to cooperation in negative social dilemmas than in positive social dilemmas has been analyzed by several social scientists working in the domain of collective action (Berejikian, 1992; Walder, 1994). Walder (1994, p.9), for instance, argued, "that loss is a more powerful motivator than gain, or that groups threatened with a loss will be more likely to protest than groups that seek proactively to achieve a gain," where protesting corresponds to cooperating. He derives this hypothesis from historical evidence and from the empirical regularity in individual decision making that actors in general are risk seeking in situations where outcomes are losses, and risk averse in situation where outcomes are gains (Kahneman and Tversky, 1979). Berejikian (1992) arrives at a similar conclusion.

Raub and Snijders (1997; Snijders and Raub, 1998) also analyzed the issue of cooperation, but in the context of indefinitely repeated two-person prisoners' dilemmas (PDs). They applied a game-theoretic analysis and came to opposite implications, of which we briefly recapitulate the main argument here. A repeated (two-person) PD is a PD game that is played repetitively in subsequent periods by the same two actors. The PD is indefinite when there is a certain fixed probability larger than zero and smaller than one that another period of the PD will be played by these two actors. Raub and Snijders claimed that these repeated PDs approximate a relevant set of collective action problems and social dilemma situations. They showed that a game-theoretic analysis of such repeated PDs results in the prediction that risk-averse individuals are *more* inclined to cooperate than risk-seeking individuals. Assuming that most individuals are risk seeking for losses and risk averse for gains, Raub and Snijders predicted larger cooperation levels in positive social dilemmas than in negative social dilemmas. They tested their hypothesis empirically in an experiment. The proportion of individuals that cooperated in the first period did not differ significantly between positive and negative dilemmas, but they did find that risk aversion promoted cooperation. Hence their results were in agreement with their prediction and falsified the hypothesis of Walder and Berejikian. In subsequent studies the game-theoretic analysis was also applied to mixed social dilemmas (van Assen, 1998; van Assen and Snijders, 2004, 2005). Van Assen showed that loss aversion, that is, the phenomenon that losses loom larger than gains (Kahneman and Tversky, 1979), promotes cooperation in mixed social dilemmas. Because experiments on individual decision making demonstrate that individuals are loss averse, he predicted that actors are even more inclined to cooperate in mixed dilemmas than in positive and negative dilemmas. Van Assen and Snijders (2004, 2005) tested this hypothesis in an experiment that involved four repeated PDs that were similar except for a shift in their outcomes. The proportion of individuals that cooperated in the first period did not differ significantly between negative, mixed, and positive dilemmas. However, there was some evidence in favor of the hypotheses that risk aversion and loss aversion promote cooperation.

In the present study, cooperation in negative, mixed, and positive PDs is investigated by reanalyzing the data of the experiment of van Assen and Snijders using a new methodology, a statistical learning model. The model developed and described in this paper has three advantages over other more traditional methods that have been used to analyze data of repeated games. Advantages are that the model allows for:

(1) an analysis that takes known empirical falsifications of standard game theoretic models into account and allows for other kinds of learning,
(2) a stronger test of the role of risk preferences in repeated PDs,
(3) a more efficient use of the available data—from about 10% to close to 100%, and
(4) can be generalized beyond the repeated PD case.

These advantages are explained subsequently.

With respect to the first advantage we start by noting that the standard game-theoretic analysis could only be applied straightforwardly to predict actors' behavior in the first period of the PD. The reason is that the standard game-theoretic analyses assume that actors are strictly forward-looking. In an analysis of repeated PDs with complete information, this implies that the behavior in the previous period should not matter, but we know from the empirical data that it does. One obvious improvement would therefore be to somehow involve what happened in previous periods in the decisions of actors in subsequent periods. This is in line with a well-known result in the empirical analyses of repeated PDs: a large number of variables affect cooperation that according to the standard game-theoretic model should not (e.g., Cain, 1998; Sally, 1995).

The observation that people not only predict but also adapt led some scientists to model actors' behavior in repeated PDs with learning models based on simple psychological assumptions on how actors choose their behavior depending on reinforcements and punishments they have received in the history of the game. In the literature on learning, 'learning' is in general not interpreted as 'discovering how to play to obtain better outcomes' but simply as 'a change in behavior'. Studies modeling behavioral dynamics in repeated games with learning models have powerfully made the case that learning models have potential for describing these dynamics (see, e.g., Chapter 6 of Camerer, 2003, and Erev and Roth, 1998, for an overview), but much remains to be desired in these models. We believe that an analysis of the data of our experiment with a learning model might also detect and explain differences in behavioral dynamics in social dilemmas differing in their range of outcomes.

Second, learning models should allow for the possibility that not only characteristics of the subjects, such as risk preferences, matter for the probability to cooperate in a repeated PD but also the "history of the game" (cf. Gautschi, 2002) matters as well. In this sense, being able to include data on behavior in later periods provides a stronger test of the hypotheses on risk preferences. It would provide strong support for the hypotheses on risk preferences if its effects surface even in later periods where the history of the game can be expected to have an important effect on behavior.

Third, another reason to employ a learning model is that traditional data analysis methods cannot handle the dependencies that exist between the responses of the actors in period 2 and beyond of the repeated PD, and between subsequently played games. That is, the probability to cooperate in a game likely depends on the history of the game and the history of playing other games. The number of possible histories grows exponentially with the number of periods that the game is played, making it impossible to devise statistical tests at a reasonable level of power. Therefore, both Raub and Snijders (1997) and van Assen and Snijders (2004, 2005) discarded all responses in period 2 and beyond of each repeated PD. They also discarded all data obtained from games played after the first games, precisely because of possible learning effects that might have occurred between games. As a result, the vast majority of the data remained unused; van Assen and Snijders (2004, 2005) even discarded more than 90% of their data because of possible learning effects intervening with effects of risk preferences. If a learning model is employed, then in principle all data can be considered in the analysis because its basic assumption is that an actor's choice is dependent on the history of choices of himself and the other actor.

To summarize, in the present study we attempt to answer the research question whether the behavioral dynamics are different for negative, mixed, and positive repeated PDs by employing a new learning model, the logistic learning model. The next two sections are devoted to learning models as a method to answer our research question. In the next section a brief sketch is provided of previous research employing learning models, and our model is situated in this line of research. Our model is explained in detail in the section Logistic Learning Model. That section also summarizes the hypotheses that we tested using our model. The experiment of van Assen and Snijders to which the model is applied is described in brief in the Experiment section. The Results section summarizes the results of applying the learning model to the data of the experiment to answer our research question. Finally, we present a conclusion and discussion, which contains an inventory of advantages of using the logistic learning model as a general tool to model behavior in any repeated game, not just the PD.

LEARNING MODELS OF BEHAVIOR

A number of different learning models have been described in the literature. The two most well-known models are the *Bush-Mosteller learning model* (1951, 1955) and the *Roth-Erev learning model* (1995; Erev and Roth, 1998, 1999). Both models are so-called reinforcement models of learning. Reinforcement models assume that the outcome of an actor's decision is 'reinforced' on the basis of the payoff an actor receives. There are many more approaches to learning in games in general and social dilemmas in particular, including evolutionary dynamics, belief learning, sophisticated anticipatory learning, experience-weighted attraction learning, imitation, direction learning, and rule learning. In the past five years or so, the number of articles focusing on learning in games using one or more of these approaches has increased dramatically. We refer the reader to Chapter 6 of Camerer's (2003) recent book on behavioral game theory for an overview of the field.

In fact, none of the models described in the literature on learning in games qualifies for our purposes. One problem of the existing models is that there is a one-to-one correspondence between most of the parameters in the model and the payoffs of the game. For example, the traditional Bush-Mosteller and Roth-Erev reinforcement models assume that an action's reinforcement is equal to the payoff it yields (Erev and Roth, 1998, 1999; Flache and Macy, 2002; Roth and Erev, 1995). Because the payoff structure is identical for the PDs used in our study, the traditional models would yield identical predictions in these PDs. Given that we precisely want to test whether behavior and learning is identical in these PDs, these traditional models are not adequate.

A second problem is that the models described in the literature do not allow for statistical tests of effects of independent variables. That is, in the present study we focus on testing whether the outcome domain (losses, gains, mixed) affects the behavioral dynamics of the repeated PD game. We want to be able to control for characteristics of the subject, the game, and the history of play. Existing models are not adequate because they do not allow incorporation of such variables. We developed a logistic learning model where the incorporation of such variables is possible, and therefore argue that this statistical model presents a major step forward in modeling behavior in repeated games. This model is described in the next section.

THE LOGISTIC LEARNING MODEL

The logistic learning model can be applied to any game with two alternatives for all the players involved in the game, but in this case we only need to focus on the two-person PD. Denoting one of the two choices of actor j in period t of the game by $Y_{jt} = 1$ and the other choice by $Y_{jt} = 0$, it is assumed that the probability of $Y_{jt} = 1$ is a logistic function of X_{jt}:

$$P(Y_{jt} = 1) = \frac{e^{X_{jt}}}{1 + e^{X_{jt}}} \tag{1}$$

Cooperation is here arbitrarily designated by '1' and defection by '0'. Hence our learning model represents a logistic regression of cooperation on X_{ji}. X_{jt} can be regarded as j's propensity to play $Y_{jt} = 1$, and is itself also a function. Function X_{jt}, called the logit in logistic regression, for a two-person game with two alternatives for each player is given by:

$$X_{jt} = \beta_0 + \sum_{r=1}^{t-1} \alpha^{t-r-1}(\beta_{00} + \beta_{10}Y_{jr} + \beta_{01}Y_{kr} + \beta_{11}Y_{jr}Y_{kr}) \tag{2}$$

There are two important differences between our model and models estimated in traditional logistic regression analysis. The first difference is the specific structure of the logit X_{jt} in our model. The second difference is that we allow for independent variables entering the

TABLE 1 The Effect of Both Actors' Choices ('0' = Defection, '1' = Cooperation) in Period r on X_{jr+1}

	$Y_{kr} = 0$	$Y_{kr} = 1$
$Y_{jr} = 0$	β_{00}	$\beta_{00} + \beta_{01}$
$Y_{jr} = 1$	$\beta_{00} + \beta_{10}$	$\beta_{00} + \beta_{10} + \beta_{10} + \beta_{11}$

model by assuming that the coefficients may depend on them, as we explain below.

In 'normal' logistic regression, the logit is a function of the independent variables. In contrast, the logit in our model is a function of behavior in previous periods and has a specific structure containing three elements. The first element is the *constant* β_0. This constant reflects the initial propensity to play 1 in the first period, or $Y_{jl} = 1$. Other reinforcement learning models, such as the Bush-Mosteller and Roth-Erev models also contain a parameter corresponding to the initial propensities to choose one of the two behavioral alternatives. However, contrary to these two models, in our model this initial propensity is not discounted. That is, the effect of the initial propensity on the total propensity to cooperate is constant across all periods. We chose not to discount β_0 to reflect the belief from personality psychology that persons are endowed with relatively stable personality traits that are not affected by a few rounds of game playing.

The second element in X_{jt} is a linear term $\beta_{00} + \beta_{10}Y_{jr} + \beta_{01}Y_{kr} + \beta_{11}Y_{jr}Y_{kr}$. This linear combination contains four parameters, one parameter for each cell in the two-by-two game. Depending on the choices of actors j and k in period r, a possibly different reinforcement or value is added to X_{jr+1}. Table 1 shows the value added to X_{jr+1} for each combination of choices.

Note that β_{00} represents the effect of both actors playing 0 in period r, while β_{01} and β_{10} represents the effects of only one actor playing 1. Parameter β_{11} can be considered as the (interaction) effect of both actors playing 1 that cannot be explained by simple addition of the effects β_{10} and β_{01}. The Bush-Mosteller and Roth-Erev models also incorporate parameters corresponding to all choice combinations.

Finally, the third element of the function X_{jt} is the geometrical discounting of the history of play in the game, represented by the sum involving a discount factor α. If $\alpha = 0$, then only choices in the last period have an effect on X_{jt}. If $\alpha = 1$, then a particular choice combination in period r exerts its effect in all future periods in the same way; that is, its effect does not diminish over the course of the game. In general, geometric discounting implies that a particular choice combination in period $r-1$ is α times as relevant for the propensity to play 1 than in period r. It is common to incorporate the history of play in a learning model by geometrically discounting actors' choices in the history of the game. For example, the Bush-Mosteller and Roth-Erev models also incorporate geometric discounting.

The second difference between traditional logistic regression analysis and our approach is where potential other independent variables (or control variables) enter the model. In normal logistic regression the logit is a direct function of the independent variables. In our model the independent variables, such as an actor's risk aversion, the payoffs of each choice combination, or other characteristics, enter Equation (2) indirectly. The independent variables are used to model the six parameters, that is, the values of these parameters are modeled as a linear combination of the variables hypothesized by the researcher to be relevant for the effect of the choice combination corresponding to the parameter. These variables can be distinguished into *subject characteristics, game characteristics, history of play characteristics*, and interactions between them. Note that the model allows that such variables can enter the equation more than once because one can model an effect of a variable on any of the different parameters.

Subject *characteristics* are variables such as sex, ethnicity, age, education, but also include variables such as risk aversion, social orientation, and intelligence. The *game characteristics* are variables involving features of the game, such as the payoffs, information of the game provided to the subjects, communication possibilities, and content of the instruction. *History of play characteristics* are variables related to the course of play in the experiment, e.g., the number of times a game is played and the period of the game.

We now apply the logistic learning model to test a number of hypotheses concerning our research question and the effects of several subject, game, and history of play characteristics that can be relevant for cooperation in the PD (Sally, 1995; Cain, 1999). The statistical or null hypotheses tested are that the value of each of the six parameters is not affected by the game characteristic PD type (negative, mixed, positive), and the subject characteristics gender, number of siblings, social orientation, risk aversion, and secondary school subjects. We also test if the history of play characteristics, 'period' and 'game number' (the number of times that the repeated PD has been played at the time or earlier), have an effect. Tests are two-sided because we have no specific expectations on the direction of a possible effect of the subject and history of play characteristics.

As far as we know, the logistic learning model cannot be fitted by logistic regression procedures in standard statistical packages without

TABLE 2 The PD in The Experiment with its Outcomes Dependent on Ego's (Row) and Alter's (Column) Choices (*C*ooperation, *D*efection). The First Number in Each Cell Represents the Outcome of Ego

	C	D
C	$10+\Delta, 10+\Delta$	$-5+\Delta, 20+\Delta$
D	$20+\Delta, -5+\Delta$	$0+\Delta, 0+\Delta$

modifications, because the logit in Equation (2) contains a product of parameters. Explicit programming of the likelihood function is necessary; we used STATA for that purpose. Our implementation allows estimates of parameters, possibly as a function of several variables (as in ordinary logistic regression), standard errors of the estimates, and tests of single parameters, sets of parameters, and complete models. The standard errors of the estimates are adjusted for the dependencies in the responses not accounted for by the variables incorporated in the model, using Huber's clustering approach (Huber, 1967). The log (pseudo-)likelihood value of a model is also reported. However, because of the dependencies in the responses, the common likelihood-ratio test cannot be performed. Tests of (sets of) parameters need to be performed by the Wald test, which are reported in the output of the analyses. The Wald test is comparable to a z-test, and to the well-known t-test of the effect of one single independent variable in multiple regression.

EXPERIMENT[1]

The Games

The subjects played four different indefinitely repeated games, all with a continuation probability equal to 0.5 that a next period is played in the PD. The outcomes of the games (T, R, P, S) were equal to $T = 20+\Delta, R = 10+\Delta, P = 0+\Delta,$ and $S = -5+\Delta,$ with the shift of outcomes $\Delta = 0$ (positive PD), $\Delta = -5$ (mixed PD), $\Delta = -10$ (mixed PD), and $\Delta = -20$ (negative PD).[2] The PD and its outcomes is presented in Table 2.

[1]See van Assen and Snijders (2004) for details of the experiment.

[2]The PD with $\Delta = 0$ and $S = -5$ was treated as a positive PD because according to standard game-theoretic predictions only the values of T, R, and P matter, and not the value of S (Raub and Snijders, 1997; van Assen, 1998).

Procedure

The experiment consisted of two parts. In the first part, risk aversion of subjects was assessed; in the second part, subjects played a number of repeated PDs (all behind the computer). The two parts were run on different days. In total, 216 subjects completed both parts of the experiment.

First Part of the Experiment

Subjects' risk aversion was assessed in two ways. First, risk aversion was assessed with a traditional method. The method required subjects to make three preference comparisons between the gamble $(20 + \Delta, \mathrm{p}, \Delta)$ and certain outcome $\mathrm{R} + \Delta$, one for each value of p equal to 1/3, 1/2, and 2/3. This procedure was repeated for each of the four values of Δ. The number of times a subject indicated a preference for $\mathrm{R} + \Delta$ over the gamble (either 0, 1, 2, or 3 times) constitutes a measure of risk aversion for the range of outcomes $\mathrm{P} + \Delta, [\mathrm{T} + \Delta]$.

Subjects' risk aversion was also assessed with the tradeoff method. The tradeoff method enabled us to estimate the subjects' utility parametrically; that is, the method generates estimates of concavity of the utility function and of loss aversion for each subject.

Finally, subjects completed a questionnaire at the end of the first part. Subjects were asked their age, sex, secondary school subjects, number of siblings, and whether they had any knowledge of game theory. The questionnaire was concluded with measurement of the social orientation of the participant. Subjects received a reward of 35 Dutch guilders (approximately 16 Euros, and at that time approximately US$ 17.5) for participation.

Second Part of the Experiment

The subjects played the PD not against another player in the lab but against a computer program. There are many complementary advantages of having subjects playing against a program instead of human subjects (see van Assen and Snijders, 2004, for a discussion). It allows experimenters to gather a maximum number of possible responses per subject per time unit, requiring as little time of as few experimenters as possible, without losing the reality of actual interaction if the program's behavior mimics human behavior very well.

The subjects played against the logistic learning model with the following specification of the linear combination:

$$X_{jt} = -0.74 + \sum_{r=1}^{t-1} \left(\frac{1}{2}\right)^{t-r-1} [-2.04 + 2.4 y_{jr} + 2.05 y_{kr}] \tag{3}$$

The value of X_{jt} was determined by substituting the history of the PD game of the subject (3). The probability p that the program cooperated was obtained by substituting the value of X_{jt} in (1). The program cooperated if the value of a draw from the uniform probability distribution was in the interval $[0, p]$, and defected if it was in the interval $[p, 1]$.

This specification of the learning model and its values of the parameters was obtained by fitting the logistic learning model to the responses in repeated PDs in a similar, previous experiment (Raub and Snijders, 1997). The model fitted the proportion of cooperative responses in the first five periods in that experiment very well.[3] Therefore, it was as if the subjects in our experiment played against a player that participated in the experiment of Raub and Snijders (1997). Subjects were told beforehand that they played against a computer program mimicking subjects' behavior in previous experiments, and that, consequently, it was as if they were playing against a randomly selected player from a previous experiment, drawn with replacement before each repeated PD. They were also told that they were not deceived in any way (which was true). Finally, note that the program's strategy is identical in all PD types. This agrees with the null hypothesis to be tested that subjects' behavior is identical in all PD types.

The subjects played the four different repeated PDs types in a random order. The subject continued to play these four repeated PDs in a random order until 40 minutes had passed, or until (s)he had played each repeated PD type 20 times.

The subject's reward in the second part of the experiment was determined by the random selection method. Their reward was a random selection of the outcome of the last period of each of the four different repeated PDs played for the first time, and the outcome of the last period of an additional PD that was included in the experiment to increase the subjects' expected payoff.

RESULTS

The results are presented for different specifications of the logistic learning model. Figure 1 gives a general idea of what the data look like. It depicts the proportion of cooperators in the first seven periods of a game, averaged across all PDs played. Periods eight or higher are not included in the figure because there are only a small number of games of this length (approximately $1/2^7 \approx 0.8\%$ of the games have

[3]The goodness of fit test comparing observed and predicted proportions yielded a chi-square value of 3.24. This result is not even significant for 1 degree of freedom, even though many more (dependent) proportions are fitted.

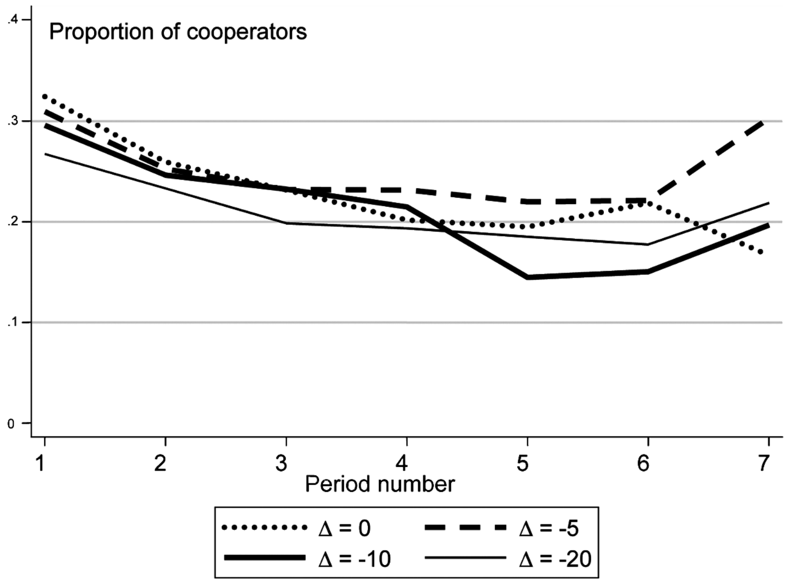

FIGURE 1 Proportion of cooperators per game in the first seven periods for the positive ($\Delta = 0$), mixed 1 ($\Delta = -5$), mixed 2 ($\Delta = -10$), and negative ($\Delta = -20$) PD.

8 or more periods). The proportion of cooperators varies between 15% and a little over 30%. Since we are interested in learning effects in the different constituent games, the graph in itself does not give much information, other than that the proportions of cooperators in each period are relatively close for the different games. Note that this does not say anything about possible effects of personal characteristics on the probability of cooperation in the repeated PD.

We first present the basic or null model with estimates of the parameters $\alpha, \beta_0, \beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}$. This model is the benchmark with which the other models are compared. We then test our main hypothesis that behavioral dynamics differ between negative, mixed, and positive repeated PDs. Finally, we briefly present the results of models that also included history of play and subject characteristics as predictors.

The Basic Model

The parameter estimates and tests whether these estimates equal zero are shown in Table 3. The log pseudo-likelihood of the model is equal to $-14{,}449.7$. The interpretation of the estimates and their tests in Table 3 is as in logistic regression. For example, $\beta_{10} = 1.25$ implies

TABLE 3 Estimates, Based on 27,253 Observations, and Tests of Parameter Estimates of the Basic Model

	Coeff.	Robust Std. Err.	Z	P > |z|
ln(α)	−0.9406	.1272	−7.39	0.000
β_{10}	1.2458	.1213	10.27	0.000
β_{01}	0.0412	.0895	0.46	0.646
β_{11}	1.4598	.1757	8.31	0.000
β_{00}	−0.9346	.0882	−10.59	0.000
β_0	−0.8177	.0945	−8.65	0.000

that unilateral cooperation of ego increases the logit of cooperation with 1.25, and multiplies the odds to cooperate with $e^{1.25} = 3.49$.

The algorithm estimates ln(α), the estimate of α then equals e to the power $\ln(\alpha) = 0.3904$. That is, on average a particular choice combination (CC, CD, DC, or DD) in period $r - 1$ is 0.39 times as relevant for the propensity to cooperate than it is in period r. On average, the proportion to cooperate in the first period is equal to $0.31\,(=e^{-0.8177}/(1+e^{-0.8177}))$. If both actors defected in the last period, then the odds to cooperate in the following period were $0.393\,(=e^{-0.9346})$ times as small ($p < 0.001$). Similarly, if only alter cooperated, then the odds in the next period were 0.409 times as small. If only ego cooperated, then the odds to cooperate in the next period increased, that is, they were 1.365 times as large. These two odds are not statistically significant, meaning that the effects of mutual defection and of unilateral cooperation of alter, are similar. Finally, if both actors cooperated, then the odds of mutual cooperation in the next period increased dramatically; they were 6.124 times as large. As expected, the estimates presented in Table 3 reveal that mutual defection increases defection and mutual cooperation increases cooperation. Surprisingly, unilateral defection only lowers the cooperation rate if alter is cooperating and ego is defecting.

Differences Between the PDs

The results of the analysis with respect to our main hypothesis, that behavior is different in negative, mixed, and positive repeated PDs, are presented in Table 4. The log pseudo-likelihood of the model, bases on 26,014 observations, is −13,588.5. The estimate of each parameter in Eq.(2) can be obtained by substituting the values of the independent variables in the equation for that parameter. For example, the estimate of β_{10} for the negative PD is equal to $1.7728 - 0.8779 = 0.8949$.

TABLE 4 Estimates, Based on 26,014 Observations, and Tests of Parameter Estimates of the Model that includes Type of PD ("POS" with $\Delta = 0$, "MIX 1" with $\Delta = -5$, "MIX 2" with $\Delta = -10$, "NEG" with $\Delta = -20$) as Predictor. "POS" is the Reference Category

	Robust			
	Coeff.	Std. Err.	Z	P > |z|
$\ln(\alpha)$				
MIX 1	−.1314	.4272	−0.31	0.758
MIX 2	.4967	.2235	2.22	0.026
NEG	.7094	.2313	3.07	0.002
cons (POS)	−1.2729	.2167	−5.88	0.000
β_{10}				
MIX 1	−.3112	.1736	−1.79	0.073
MIX 2	−.6866	.1676	−4.10	0.000
NEG	−.8779	.1776	−4.94	0.000
cons (POS)	1.7728	.1740	10.19	0.000
β_{01}				
MIX 1	.2807	.1847	1.52	0.128
MIX 2	−.1019	.1949	−0.52	0.601
NEG	−.1504	.1838	−0.82	0.413
cons (POS)	.0865	.1712	0.51	0.614
β_{11}				
MIX 1	−.3254	.3090	−1.05	0.292
MIX 2	−.4193	.3315	−1.26	0.206
NEG	−.5367	.3516	−1.53	0.127
cons (POS)	1.7158	.2760	6.22	0.000
β_{00}				
MIX 1	.1627	.1245	1.31	0.191
MIX 2	.5154	.1168	4.41	0.000
NEG	.7614	.1233	6.17	0.000
cons (POS)	−1.3340	.1361	−9.80	0.000
β_0				
MIX 1	−.0670	.0453	−1.48	0.139
MIX 2	−.1171	.0633	−1.85	0.064
NEG	−.2707	.0946	−2.86	0.004
cons (POS)	−.7370	.1124	−6.56	0.000

The results in Table 4 show that there are some clear differences in behavior in the four PDs. The results of Table 4 can be briefly summarized as follows: the proportion of cooperation in the first period (β_{00}) is higher in the positive PD than in the negative PD (0.324 versus 0.267, $p = 0.004$), and not statistically different from the mixed PDs. The effects of unilateral cooperation of alter (β_{01}) and of mutual cooperation (β_{11}) on cooperation are not statistically different for the four PDs. The shadow of the past was also larger in these two

PDs: 0.46 for the mixed 2 PD ($\Delta = -10$) and 0.569 for the negative PD, versus 0.28 for the positive PD and 0.246 for the other mixed PD. Finally, the effect of mutual defection (β_{00}) on cooperation is larger (i.e., less negative), and the effect of unilateral cooperation of ego (β_{10}) on cooperation is smaller (i.e., less positive) for the two PDs with the largest negative shift ($\Delta = -10$ and $\Delta = -20$) of outcomes ($p < 0.001$).

Other Models

The previous analyses are the ones that are most relevant to our main hypothesis concerning our research involving the effect of the circumstances (bad, mixed, good) on cooperation in social dilemmas. However, to illustrate the applicability and versatility of the proposed statistical logistic learning model, in this section the results of analyses that incorporate additional predictors are briefly discussed.

First, an analysis with only risk aversion as assessed with the traditional method demonstrated that the differences in behavior between the four PDs cannot be explained by risk aversion if we consider all the data simultaneously (instead of just the first period). Only a significant positive effect of risk aversion on mutual defection was observed ($p = 0.002$). That is, on average, the more risk averse, the more likely it is that ego will defect after mutual defection in the previous period.

In addition, we ran three analyses, all based on 25,276 observations, that include history of play and subject characteristics as predictors of the parameters $\alpha, \beta_0, \beta_{00}, \beta_{01}, \beta_{10}, \beta_{11}$, additional to the predictors corresponding to game type as in Table 4. Firstly, an analysis was run with game type (MIX 1, MIX 2, Neg), with period (Period), and with interactions of game type and period (MIX1xPer, MIX2xPer, NegxPer) as predictors (log pseudo-likelihood $= -13{,}160.3$). With Period we tested whether the effect of 'DD', DC', 'CD', 'CC' on cooperation changed over periods. Secondly, an analysis was run with game number (Game_nr) together with the predictors of the first analysis (log pseudo-likelihood $= -13{,}131.5$). With game number we tested whether the initial propensity to cooperate, discounting, and the effect of 'DD', DC', 'CD', 'CC' on cooperation changed after playing more PD games. The first and second analysis only added history of play variables to the basic model. In the third and final analysis subject characteristics were included together with all the predictors of the second analysis (log pseudo-likelihood $= -12{,}923.5$). The subject characteristics included are gender (Gender), number of siblings (Siblings), number of science subjects in secondary school, like math, physics, chemistry and biology (Beta), risk aversion (Risk), social orientation ('No_so' indicating whether the subject did give consistent

TABLE 5 Results of Wald Tests Whether a Predictor Affects Behavior in the Repeated PDs. The Tests are Based on 25,276 Observations

Predictor	Wald	Df	p-value
Game type	68.72	18	<0.001
Game type, Period, Game type × Period	151.70	34	<0.001
Game_nr	24.85	6	<0.001
Gender	16.29	6	0.012
Siblings	10.17	6	0.118
Beta	5.05	6	0.537
Risk	12.53	6	0.051
So	18.99	6	0.004
No_so	24.85	6	<0.001

answers to the social orientation questions in the questionnaire, 'So' as a measure of social orientation if the subject did give consistent answers). The fit of the data improved significantly after each analysis. Only the results of the third analysis are reported. Table 5 provides the results of Wald tests on whether a set of predictors affects behavior in the repeated PDs. Table 6 presents the estimates and tests of single parameters in the model.

The first two lines of Table 5 show that Game type and the history of play characteristics, Period and interactions of Game type and Period, do affect behavior in the repeated PDs. The effects of Game type we have already described above. From Table 6 we learn that the effect of unilateral cooperation of alter increased (effect of Period on β_{01}) over periods, and the effect of unilateral cooperation of ego increased over periods for the two games with the worst outcomes compared to the positive PD (effect of MIX2xPer and NegxPer on β_{10}). Behavior in the repeated PDs was also affected by Game_nr, that is, the number of PDs a subject had already played. The more PDs subjects had played, the higher the initial propensity to cooperate (β_0), the more negative the effect of mutual defection (β_{00}) and mutual cooperation (β_{11}), and the more positive the effect of unilateral cooperation (β_{10} and β_{01}) on cooperation. Hence, subjects started more frequently with cooperation but learned over rounds of play that after mutual defection there was not much hope to obtain mutual cooperation. To derive what these results imply for unilateral and mutual cooperation, the parameter estimates must be substituted in the expressions for unilateral and mutual cooperation in Table 1. Tests of these expressions reveal that the effect of unilateral and mutual cooperation did not change after playing more PDs.[4]

[4]That is, each of β_{00} and β_{11} is cancelled out by each of β_{10} and β_{01}.

TABLE 6 Estimates, Based on 25,276 Observations, and Tests of Parameter Estimates of the Model that includes Type of PD, Subject Characteristics (Gender, Siblings, Social Orientation ('no_so' and 'so'), Beta, Risk Preferences), and History of Play Characteristics (Period, Game Type × Period Interactions, Game Number) as Predictors.

	Robust			
	Coeff.	Std. Err.	Z	P > |z|
$\ln(\alpha)$				
MIX 1	.1435	.5257	0.27	0.785
MIX 2	.4712	.4052	1.16	0.245
NEG	.5369	.3570	1.50	0.133
Game_nr	.0001	.0066	0.02	0.987
Gender	−.1429	.3118	−0.46	0.647
Siblings	.1947	.1082	1.80	0.072
So	.1165	.0680	1.71	0.087
No_so	.0273	.4174	0.07	0.948
Beta	−.1130	.1076	−1.05	0.294
Risk	.0798	.1373	0.58	0.561
Cons (POS)	−1.5761	.5597	−2.82	0.005
β_{10}				
MIX 1	−.6921	.4206	−1.65	0.100
MIX 2	−1.6278	.4251	−3.83	0.000
NEG	−2.0248	.4452	−4.55	0.000
Period	−.2315	.1200	−1.93	0.054
MIX1xPer	.1267	.1414	0.90	0.370
MIX2xPer	.3111	.14344	2.17	0.030
NegxPer	.3904	.14153	2.76	0.006
Game_nr	.0114	.0035	3.27	0.001
Gender	.7080	.2531	2.80	0.005
Siblings	−.01228	.1347	−0.09	0.927
So	−.0987	.1259	−0.78	0.433
No_so	−.3794	.3898	−0.97	0.330
Beta	.0624	.0906	0.69	0.491
Risk	−.1275	.1225	−1.04	0.298
Cons (POS)	1.9652	.7153	2.75	0.006
β_{01}				
MIX 1	.8137	.5191	1.57	0.117
MIX 2	.0385	.5110	0.08	0.940
NEG	−.1423	.5229	−0.27	0.785
Period	.28197	.1356	2.08	0.038
MIX1xPer	−.2028	.1690	−1.20	0.230
MIX2xPer	−.0994	.1627	−0.61	0.541
NEGxPer	−.0142	.1607	− 0.09	0.930
Game_nr	.0117	.0036	3.21	0.001
Gender	.5201	.1918	2.71	0.007
Siblings	.0749	.0794	0.94	0.346

(*Continued*)

TABLE 6 Continued

	Robust			
	Coeff.	Std. Err.	Z	P > |z|
So	.0506	.0804	0.63	0.529
No_so	.0762	.2816	0.27	0.787
Beta	.0906	.0686	1.32	0.187
Risk	−.1458	.1061	−1.37	0.170
Cons (POS)	−1.5423	.4880	−3.16	0.002
β_{11}				
MIX 1	−.4111	.8179	−0.50	0.615
MIX 2	−.0600	.8556	−0.07	0.944
NEG	.1336	.8386	0.16	0.873
Period	−.1216	.2378	−0.51	0.609
MIX1xPer	.0380	.2794	0.14	0.892
MIX2xPer	−.0714	.2920	−0.24	0.807
NegxPer	−.1895	.2680	−0.71	0.479
Game_nr	−.0181	.0056	−3.24	0.001
Gender	−1.2088	.3667	−3.30	0.001
Siblings	−.2984	.1795	−1.66	0.096
So	.1583	.1271	1.25	0.213
No_so	.7904	.5269	1.50	0.134
Beta	.0505	.1256	0.40	0.687
Risk	.1648	.1780	0.93	0.355
Cons (POS)	2.9364	1.0451	2.81	0.005
β_{00}				
MIX 1	.1065	.2594	0.41	0.682
MIX 2	.9182	.2702	3.40	0.001
NEG	1.009	.2628	3.84	0.000
Period	.0157	.0651	0.24	0.809
MIX1xPer	.0348	.0804	0.43	0.666
MIX2xPer	−.0947	.0829	−1.14	0.253
NEGxPer	−.0913	.0735	−1.24	0.214
Game_nr	−.0101	.0025	−3.96	0.000
Gender	−.2906	.1644	−1.77	0.077
Siblings	−.0194	.0687	−0.28	0.778
So	.08277	.05918	1.40	0.162
No_so	−.0998	.2385	−0.42	0.676
Beta	−.0728	.0590	−1.23	0.217
Risk	.2776	.0804	3.45	0.001
Cons (POS)	−1.2884	.3695	−3.49	0.000
β_0				
MIX 1	−.0886	.04956	−1.79	0.074
MIX 2	−.2062	.0696	−2.96	0.003
NEG	−.2663	.0995	−2.68	0.007
Game_nr	.0041	.0020	2.03	0.042
Gender	.1634	.2051	0.80	0.426

(*Continued*)

TABLE 6 Continued

	Robust			
	Coeff.	Std. Err.	Z	P > |z|
Siblings	−.0140	.1070	−0.13	0.896
So	.0709	.0732	0.97	0.333
No_so	.6886	.2712	2.54	0.011
Beta	−.0052	.0643	−0.08	0.936
Risk	−.1806	.0937	−1.93	0.054
Cons (POS)	.9100	.3743	−2.43	0.015

Further inspection of Table 5 and Table 6 reveals that there was an effect of the subject characteristics risk aversion, gender and social orientation on subjects' behavior in repeated PDs. The effect of risk aversion we have already described above. Women tend to have a larger value of $\beta_{10}(p = 0.005)$ and $\beta_{01}(p = 0.007)$, but a smaller value of $\beta_{00}(p = 0.077)$ and $\beta_{11}(p = 0.001)$. However, after substituting the parameter estimates in the expression of Table 1, tests revealed that the effects of unilateral and mutual cooperation on cooperation in the next round were not significantly different for men and women.[4] Finally, although the overall tests of the effects of social orientation (So and No_so) were significant, only one single parameter estimate was significantly different from zero. Subjects who did not give consistent answers to the social orientation questions were more inclined to cooperate in the first round of the PD ($p = 0.011$). We have no meaningful interpretation of this result.

CONCLUSION AND DISCUSSION

Our main concern in the present and in previous studies was to establish whether there are behavioral differences in negative, mixed, and positive dilemmas, and how such differences can be explained. The logistic learning model described here was in the first place developed to solve the problems we encountered in previous studies to investigate that concern. Game theoretic tools did not help us much in deriving predictions of behavior in periods following the first period. As a consequence, most of the data could not be incorporated in the data analysis because standard logistic regression analyses cannot adequately handle the dependencies between the responses in a repeated game. A solution to both issues was to model the data with a learning model. Existing learning models like the Bush-Mosteller and Roth-Erev models did not suffice for two reasons: application of them did not take characteristics of the game, of the history of play, and of subject characteristics into account. More importantly, it is not possible to test the effects of incorporated variables statistically. We therefore developed a learning model in which predictors (game, history of play, and subject characteristics) can be included to model and test the learning parameters. The learning model is based on logistic regression, and can be considered as a general tool for the data analysis of any 2×2 game based on the general assumption of geometric discounting. Hence, the scientific value of the logistic learning model is wider than its application to repeated PDs in the present study.

As an example application of the model, the data of the experiment of van Assen and Snijders (2004, 2005) were reanalyzed with this logistic learning model in order to (i) investigate our main research question whether behavior is different in negative, mixed, and positive dilemmas, and (ii) to discover if any of the subject, game, and history of play characteristics have an effect on behavior. We indeed found that subjects' behavior is different in the four PD types. We observed the following differences in behavior in the four PDs. First, there was more cooperation in the first period in positive PDs than in the negative PDs. This result corroborates the findings of Raub and Snijders (1997) and goes against the common intuition that it is more difficult to establish cooperation "under negative circumstances" (Berejikian, 1992; Walder, 2000). Second, in the two PDs with the lowest outcomes ($\Delta = -10$ and $\Delta = -20$) (i) the effect of previous on current behavior was stronger, (ii) the effect of ego's unilateral cooperation on cooperation was smaller, and (iii) the effect of mutual defection on cooperation was higher. Finally, some effects of risk aversion, gender, and history of play or period were observed. These effects were primarily estimated and tested to demonstrate the applicability and versatility of the statistical logistic learning model. That is, we did not have a theory concerning the effects of any of the subject characteristics on subject behavior after period 1.

The logistic learning model is employed in the present study as an alternative and flexible way to analyze the dependent data of repeated 2×2 games. We plan to extend this model to be able to tackle behavior in a wider class of games with more than two choice alternatives for all, possibly more than two, actors. A generalization of the learning model to repeated games with more than two choice alternatives is straightforward by combining the logic of the learning model with multinomial logistic regression.

Interestingly, the learning model can also be used to aid the researcher in the construction of an experiment, either as a manipulation variable or as a method to increase the efficiency of the experiment. Let us start with the efficiency. As we have shown in this experiment, the learning model allows an easy way for experimenters to let subjects play against a "randomly chosen other person from previous experiments." By fitting a learning model on previous data, one can honestly say that a random drawing is made from the previous subject pool. Of course, to be able to use the model as a player, one must be guaranteed that the program provides an accurate and realistic description of the play of the population of subjects of interest. By first carrying out a pilot study and fitting the model to the data of the pilot one can attempt to realize that.

The learning model can also be used as a manipulation tool in experiments. For example, one could let subjects in different conditions play against different populations of "players" that are represented by other values of the parameters in the logistic learning model.

REFERENCES

Berejikian, J. (1992). Revolutionary collective action and the agent-structure problem. *American Political Science Review*, *86*, 647–657.

Bush, R. R. & Mosteller, F. (1951). A mathematical model for simple learning. *Psychological Review*, *58*, 313–323.

Bush, R. R. & Mosteller, F. (1955). *Stochastic Models for Learning*, New York: Wiley.

Cain, M. J. G. (1998). An experimental investigation of motives and information in the prisoner's dilemma game. *Advances in Group Processes*, *15*, 133–160.

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction*, Princeton, NJ: Princeton University Press.

Erev, I. & Roth, A. E. (1998). Predicting how people play games: reinforcement learning in experimental games with unique, mixed strategy equilibria. *The American Economic Review*, *88*, 848–881.

Erev, I. & Roth, A. E. (1999). On the role of reinforcement learning in experimental games: the cognitive game-theoretic approach. In D. V. Budescu, I. Erev, & R. Zwick, (eds.), *Games and Human Behavior: Essays in Honor of Amnon Rapoport*. (pp. 53–77), Mahwah, NJ: Lawerence Erlbaum Associates.

Flache, A. & Macy, M. W. (2002). Stochastic collusion and the power law of learning: A general reinforcement learning model of cooperation. *Journal of Conflict Resolution*, *46*, 629–653.

Gautschi, T. (2002). *Trust and Exchange: Effects of Temporal Embeddedness and Network Embeddedness on Providing and Dividing a Surplus*, Amsterdam: Thela thesis.

Huber, P. J. (1967). The behavior of maximum likelihood estimates under non-standard conditions. *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, *1*, 221–233.

Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*, 263–291.

Raub, W. & Snijders, C. (1997). Gains, losses, and cooperation in social dilemmas and collective action: the effects of risk preferences. *Journal of Mathematical Sociology*, *22*, 263–291.

Roth, A. E. & Erev, I. (1995). Learning in extensive-from games: experimental data and simple dynamic models in the intermediate term. *Games and Economic Behavior*, *8*, 164–212.

Sally, D. (1995). Conversation and cooperation in social dilemmas: A meta-analysis of experiments from 1958 to 1992. *Rationaliy and Society*, *7*, 58–92.

Snijders, C. & Raub, W. (1998). Revolution and risk. Pardoxical consequences of risk aversion in interdependent situations. *Rationality and Society*, *10*, 405–425.

Van Assen, M. A. L. M. (1998). Effects of individual decision theory assumptions on predictions of cooperation in social dilemmas. *Journal of Mathematical Sociology*, *23*, 143–153.

Van Assen, M. A. L. M. & Snijders, C. (2004). Effects of risk preferences in social dilemmas: a game-theoretical analysis and evidence from two experiments. In Ramzi, Suleiman, Budescu, Fischer, & Messick, (eds.), *Contemporary Psychological Research on Social Dilemmas*, Cambridge: Cambridge University Press.

Van Assen, M. A. L. M. & Snijders, C. (2005). The effect of nonlinear utility on behavior in repeated prisoner's dilemmas. *Submitted for publication*.

Walder, A. G. (2000). Implications of loss avoidance for theories of social movements. *Hong Kong Journal of Sociology*, *1*, 83–102.